

- | | | | |
|-----|---|--|-------------------------------|
| RA | Merkulov G., Milshina N.V., Mobarry C., Morris J.J., Moshrefi A., Mount S.M., Moy M., Murphy B., Murzny D.M., Nelson D.L., Nelson D.R., Nixon K.A., Nixon M., Nussekern D.R., Pacleb J.M., Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G., Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H., Shue B.C., Siden-Kiamos I., Simperton M.P., Smith T., Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., Wang X., Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang J., Williams S.M., Woodage T., Worley K.C., Wu D., Weissbach J., Ye J., Yeh R.-F., Zaveri J.S., Zhang G., Zhao Q., Zheng L., Zheng X.H., Zhong N., Zhou X., Zhu S., Zhu X., Smith H.O., Gibbs R.A., Myers E.W., Rubin G.E., Venter J.C., "The genome sequence of <i>Drosophila melanogaster</i> ," <i>Science</i> 287:2185-2195(2000). | Query Match
Matches 28;
Conservative
Mismatches 3;
Gaps 0; | Score 118; DB 2; Length 2347; |
| [2] | NUCLEOTIDE SEQUENCE;
MEDLINE=22426065; PubMed=12537568; | RN | RX |
| RA | Celniker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A., Patel S., Adams M., Champé M., Dugan S.P., Frise E., Halpern A., George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R., Pacleb J.M., Park S., Pfeiffer B.D., Richardson S., Sodergren E.J., Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C., Weissbach J., "Finishing a whole-genome shotgun: release 3 of the <i>Drosophila melanogaster</i> euchromatic genome sequence," <i>Genome Biol.</i> 3:RESEARCH0079-RESEARCH0079(2002). | RN | RX |
| [3] | NUCLEOTIDE SEQUENCE;
MEDLINE=22426070; PubMed=12537573; | RN | RX |
| RA | Kaminker J.S., Bergman C.M., Krommiller B., Carlson J.W., Svirskas R., Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M., Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C., Weissbach J., "The transposable elements of the <i>Drosophila melanogaster</i> euchromatin: a genomic perspective," <i>Genome Biol.</i> 3:RESEARCH0084-1-RESEARCH0084-20(2002). | RN | RX |
| [4] | NUCLEOTIDE SEQUENCE;
MEDLINE=22426079; PubMed=12537572; | RN | RX |
| RA | Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S., Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E., Smith C.D., Tupy J.L., Whitefield B.J., Bayraktaroglu L., Berman B.P., Bettencourt B.R., Kambour K.S.E., Lewis S.E., Rubin G.M., Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q., Stapleton M., Yanada C., Ashburner M., Rubin G.M., Lewis S.E., "Annotation of the <i>Drosophila melanogaster</i> euchromatic genome: a systematic review," <i>Genome Biol.</i> 3:RESEARCH0083-1-RESEARCH0083-22(2002). | RN | RX |
| [5] | NUCLEOTIDE SEQUENCE;
BERKELEY DROSOPHILA GENOME PROJECT;
Celniker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R., Yu C., Rubin G., "Drosophila melanogaster release 4 sequence," Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases. | RN | RX |
| [6] | NUCLEOTIDE SEQUENCE;
Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases. | RN | RX |
| [7] | -1- INTERACTION:
Q9VH0_0da2j3; NbExp=1; IntAct:EBI-112777, EBI-109247;
Q9VZP3_cyc2j; NbExp=1; IntAct:EBI-112777, EBI-187013;
P37236_Ptq; NbExp=1; IntAct:EBI-112777, EBI-108126;
P4325_Nca; NbExp=1; IntAct:EBI-112777, EBI-149846; | RN | RX |
| CC | EMBL; AE003698; AAC14338.1; -; Genomic_DNA.
IntAct; Q8TWH9; - | DR | RX |
| CC | Ensembl; CG151; <i>Drosophila melanogaster</i> . | DR | RX |
| CC | FlyBase; FBgn0038108; CG7518.
InterPro; IPR001005; Myb_DNA_binding_protein; | DR | RX |
| CC | PROSITE; 2347_AA; 257013_MW; 23B8F5C5FFCAEA64_CRC64; | DR | RX |
| CC | SQ | RT | RX |
| CC | "The genome sequence of <i>Drosophila melanogaster</i> ," Science 287:2185-2195(2000). | RL | RX |
| [2] | NUCLEOTIDE SEQUENCE;
MEDLINE=22426066; PubMed=12537568; | RN | RX |
| RA | Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A., Celtniker S.E., Adams M., Champé M., Dugan S.P., Frise E., Hodgeson A., George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R., Pacioli J.M., Park S., Pfeiffer B.D., Richardson S., Soderberg E.J., Svirskas R., Tabor P.E., Wan K., "The genome sequence of <i>Drosophila melanogaster</i> ," Science 287:2185-2195(2000). | RN | RX |

RESULT	8	Q871E9_NEUCR	NEUCR PRELIMINARY;	PRT;	1240 AA.
Q871E9		ID			
AC	Q871E9;				
DT	01-JUN-2003	(TREMBL)re1.	24,	Created)	
DT	01-JUN-2003	(TREMBL)re1.	24,	Last sequence update)	
DT	01-MAR-2004	(TREMBL)re1.	26,	Last annotation update)	
DE		Hypothetical protein	B7H23.	.320.	
GN	Name=B7H23..320;				
OS	Neurospora crassa				
OC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;				
OC	Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.				
OX					
RN	[1]				
RP	NUCLEOTIDE SEQUENCE.				
RA	Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,				
RA	Nakatura G., Mewes H.W., Mannhaupt G.,				
RL	Submitted (MAR-2003) to the EMBL/GenBank/DDBJ databases.				
RN	[2]				
RN	NUCLEOTIDE SEQUENCE.				
RA	German Neurospora Genome Project;				
RL	Submitted (MAR-2003) to the EMBL/GenBank/DDBJ databases.				
DR	EMBL: BX29426; CAD71057.1; -; Genomic DNA.				
DR	InterPro: IPR001611; IPR01611; -; Genomic DNA.				
DR	PFAM: PF00560; LRR_1.1.				
DR	PRINTS: PR00019; DEURCHRPT.				
KW	Hypothetical protein; Leucine-rich repeat; Repeat.				
SEQUENCE	1240 AA;	134655 MW;	37660713FBA9CD58	CRC64;	
Query Match	52.2%	Score 109	DB 2;	Length 1240;	
Best Local Similarity	57.9%	Pred. No. 0.335;			Gaps
Matches 22;	Conservative 10;	Mismatches 6;	Indels 0;		

RESULT 10	Q7SZ95_XENIA	PRELIMINARY;	PRT;	143 AA.
ID	Q7SZ95;			
DT	01-OCT-2003	(T-EMBLrel. 25, Created 01-MAR-2004 (T-EMBLrel. 25, Last annotation update)		
DE	Hypothetical protein (Fragment).			
OS	Xenopus laevis (African clawed frog).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;			
OC	Xenopodinae; Xenopus; Xenopus.			
OX	NCBI_TaxID=8355			
RN	[1]	NUCLEOTIDE SEQUENCE.		
RP	RC	NUCLEOTIDE SEQUENCE. TISSUE=Whole;		
RX	RX	MBIDLINE=22388257;	PubMed=12477932;	DOI=10.1073/pnas.242603899;
RA	RA	Strausberg R.E., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.		
RA	RA	Klaussner R.D., Collins F.S., Wagner L., Shenmier C.M., Schaefer C.P., Bhat N.K.,		
RA	RA	Altschul S.F., Zeeberg B., Buetow K.H., Hsieh F., Hong L., Rubin G.M.,		
RA	RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,		
RA	RA	Diatrichenko L., Marsusina K., Farmer A.A., Casavant T.L., Scheetz T.		
RA	RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Prange C.		
RA	RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Mullahay S.		
RA	RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Gunaratne P.H.		
RA	RA	Bosak S.A., McEvany P.J., McKernan K.J., Malek J.A., Gay L.J., Hulyk S.		
RA	RA	Richards S., Worley K.C., Hale S., Garcia A.M., Villalon D.K., Muzy N.D.M., Sodergren E.J., Lu X., Gibbs R.A., Pahey J., Heiton E., Kettman M., Madan A., Rodrigues S., Sanchez		
RA	RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakeslee R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Rodriguez R.A.		

OC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;	PRT;	1500 AA.
OC	Sordariomycetes incertae sedis; Magnaporthe.		
NCBI_TaxID=242507;			
RN			
	NUCLEOTIDE SEQUENCE.		
	STRAIN=70-15;		
	RC		
	Birren B., Allen N., Allen T., An P., Anderson M., Anderson S., Arachchi H., Armbruster J., Bachantsang P., Baldwin J., Barry A., Borowsky M., Boukhalter B., Bloom T., Biye J., Boguslavskiy L., Calvo S., Camarat J., Campo K., Chang J., Cheshatsang Y., Citteroen M., Collymore A., Conidine T., Cook A., Cooke P., Corum B., Cuomo C., David R., Dawe T., Degray S., Dodge S., Dooley K., Dorje P., Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R., Erickson J., Farina A., Faro S., Fischer P., Galagan J., Gearin G., Ginerre S., Girkir A., Goettte A., Graham J., Grandbois E., Gyaltzen K., Hafez N., Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H., Honan T., Horn A., Houde N., Hughes L., Husby E., Iliev I., Jaffe D., Jones C., Kamal M., Kamyszellis M., Karlsson E., Keil J., Kleu A., Kiser A., Kodira C., Labutti K., Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T., Lindblad-ton K., Liu X., Lokyitsang T., Lucien O., Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J., Manning J., Marabelli R., Maru K., Mathews C., Mauceli E., McCarthy M., Mcdonough S., McGhee T., Melidim J., Meneus L., Mozes J., Mulrain L., Munhwa T., Munga V., Moru K., Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C., Norbu N., O'donnell P., Okwo O., O'leary P., Omran S., Parker S., Perrin D., Phunkhang P., Pigani B., Purcell S., Rachupka T., Ranasamy U., Rameau R., Ray V., Raymond C., Reita R., Ries C., Rodriguez J., Rogers J., Rogov P., Rutman M., Schupbach R., Seaman C., Settipalli S., Sharpe T., Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Songuez C., Spencer B., Stalker J., Stone C., Stone S., Stubbs M., Stavropoulos S., Stretton K., Tenzing P., Tesfaye S., Theodore J., Tholoutsang Y., Topham K., Tovey S., Teamla T., Tsomo N., Tuvali D., Vassilieff H., Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T., Wangdi T., Whittaker C., Wilkinson J., Wu Y., Wyman D., Yadav S., Zimmer A., Zody M., Zander B.;		
	"The Genome sequence of Magnaporthe grisea.";		
RN	Submitted (OCT-2003) to the EMBL/GenBank/DDBJ databases.		
RN			
	NUCLEOTIDE SEQUENCE.		
	STRAIN=70-15;		
	RC		
	Dean R., Mitchell T., Brown D., Pan H., Thon M.; Submitted (OCT-2003) to the EMBL/GenBank/DDBJ databases.		
	CC		
	CC CAUTION: The sequence shown here is derived from an EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is preliminary data.		
	DR		
	Hypothetical protein.		
	SQ		
	SEQUENCE 823 AA: 83110 MW;	CAC6C9027312BBBA CRC64;	
	Query Match	Score 105.5; DB 2;	Length 823;
	Best Local Similarity	59.1%; Pred. No. 0.47;	
	Matches 26; Conservative 7; Mismatches 8; Indels 3; Gaps 1;		
	KW		
	CC Preliminary data.		
	DR		
	EMBL; ARCU0100400; EAA54112.1; - ; Genomic_DNA.		
	RL		
	Submitted (OCT-2003) to the EMBL/GenBank/DDBJ databases.		
	CC		
	CC CAUTION: The sequence shown here is derived from an EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is preliminary data.		
	DR		
	Hypothetical protein.		
	SQ		
	SEQUENCE 823 AA: 83110 MW;	CAC6C9027312BBBA CRC64;	
	Query Match	Score 105.5; DB 2;	Length 823;
	Best Local Similarity	59.1%; Pred. No. 0.47;	
	Matches 26; Conservative 7; Mismatches 8; Indels 3; Gaps 1;		
	KW		
	CC Preliminary data.		
	DR		
	EMBL; ARCU0100400; EAA54112.1; - ; Genomic_DNA.		
	RL		
	Submitted (OCT-2003) to the EMBL/GenBank/DDBJ databases.		
	CC		
	CC CAUTION: The sequence shown here is derived from an EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is preliminary data.		
	DR		
	Hypothetical protein.		
	SQ		
	SEQUENCE 823 AA: 83110 MW;	CAC6C9027312BBBA CRC64;	
	Query Match	Score 105.5; DB 2;	Length 823;
	Best Local Similarity	59.1%; Pred. No. 0.47;	
	Matches 26; Conservative 7; Mismatches 8; Indels 3; Gaps 1;		
	KW		
	CC Preliminary data.		
	DR		
	EMBL; ARCU0100400; EAA54112.1; - ; Genomic_DNA.		
	RL		
	Submitted (OCT-2003) to the EMBL/GenBank/DDBJ databases.		
	CC		
	CC CAUTION: The sequence shown here is derived from an EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is preliminary data.		
	DR		
	Hypothetical protein.		
	SQ		
	SEQUENCE 823 AA: 83110 MW;	CAC6C9027312BBBA CRC64;	
	Query Match	Score 105.5; DB 2;	Length 823;
	Best Local Similarity	59.1%; Pred. No. 0.47;	
	Matches 26; Conservative 7; Mismatches 8; Indels 3; Gaps 1;		
	KW		
	CC Preliminary data.		
	DR		
	EMBL; ARCU0100400; EAA54112.1; - ; Genomic_DNA.		
	RL		
	Submitted (OCT-2003) to the EMBL/GenBank/DDBJ databases.		
	CC		
	CC CAUTION: The sequence shown here is derived from an EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is preliminary data.		
	DR		
	Hypothetical protein.		
	SQ		
	SEQUENCE 823 AA: 83110 MW;	CAC6C9027312BBBA CRC64;	
	Query Match	Score 105.5; DB 2;	Length 823;
	Best Local Similarity	59.1%; Pred. No. 0.47;	
	Matches 26; Conservative 7; Mismatches 8; Indels 3; Gaps 1;		
	KW		
	CC Preliminary data.		
	DR		
	EMBL; ARCU0100400; EAA54112.1; - ; Genomic_DNA.		
	RL		
	Submitted (OCT-2003) to the EMBL/GenBank/DDBJ databases.		
	CC		
	CC CAUTION: The sequence shown here is derived from an EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is preliminary data.		
	DR		
	Hypothetical protein.		
	SQ		
	SEQUENCE 823 AA: 83110 MW;	CAC6C9027312BBBA CRC64;	
	Query Match	Score 105.5; DB 2;	Length 823;
	Best Local Similarity	59.1%; Pred. No. 0.47;	
	Matches 26; Conservative 7; Mismatches 8; Indels 3; Gaps 1;		
	KW		
	CC Preliminary data.		
	DR		
	EMBL; ARCU0100400; EAA54112.1; - ; Genomic_DNA.		
	RL		
	Submitted (OCT-2003) to the EMBL/GenBank/DDBJ databases.		
	CC		
	CC CAUTION: The sequence shown here is derived from an EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is preliminary data.		
	DR		
	Hypothetical protein.		
	SQ		
	SEQUENCE 823 AA: 83110 MW;	CAC6C9027312BBBA CRC64;	
	Query Match	Score 105.5; DB 2;	Length 823;
	Best Local Similarity	59.1%; Pred. No. 0.47;	
	Matches 26; Conservative 7; Mismatches 8; Indels 3; Gaps 1;		
	KW		
	CC Preliminary data.		
	DR		
	EMBL; ARCU0100400; EAA54112.1; - ; Genomic_DNA.		
	RL		
	Submitted (OCT-2003) to the EMBL/GenBank/DDBJ databases.		
	CC		
	CC CAUTION: The sequence shown here is derived from an EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is preliminary data.		
	DR		
	Hypothetical protein.		
	SQ		
	SEQUENCE 823 AA: 83110 MW;	CAC6C9027312BBBA CRC64;	
	Query Match	Score 105.5; DB 2;	Length 823;
	Best Local Similarity	59.1%; Pred. No. 0.47;	
	Matches 26; Conservative 7; Mismatches 8; Indels 3; Gaps 1;		
	KW		
	CC Preliminary data.		
	DR		
	EMBL; ARCU0100400; EAA54112.1; - ; Genomic_DNA.		
	RL		
	Submitted (OCT-2003) to the EMBL/GenBank/DDBJ databases.		
	CC		
	CC CAUTION: The sequence shown here is derived from an EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is preliminary data.		
	DR		
	Hypothetical protein.		
	SQ		
	SEQUENCE 823 AA: 83110 MW;	CAC6C9027312BBBA CRC64;	
	Query Match	Score 105.5; DB 2;	Length 823;
	Best Local Similarity	59.1%; Pred. No. 0.47;	
	Matches 26; Conservative 7; Mismatches 8; Indels 3; Gaps 1;		
	KW		
	CC Preliminary data.		
	DR		
	EMBL; ARCU0100400; EAA54112.1; - ; Genomic_DNA.		
	RL		
	Submitted (OCT-2003) to the EMBL/GenBank/DDBJ databases.		
	CC		
	CC CAUTION: The sequence shown here is derived from an EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is preliminary data.		
	DR		
	Hypothetical protein.		
	SQ		
	SEQUENCE 823 AA: 83110 MW;	CAC6C9027312BBBA CRC64;	
	Query Match	Score 105.5; DB 2;	Length 823;
	Best Local Similarity	59.1%; Pred. No. 0.47;	
	Matches 26; Conservative 7; Mismatches 8; Indels 3; Gaps 1;		
	KW		
	CC Preliminary data.		
	DR		
	EMBL; ARCU0100400; EAA54112.1; - ; Genomic_DNA.		
	RL		
	Submitted (OCT-2003) to the EMBL/GenBank/DDBJ databases.		
	CC		
	CC CAUTION: The sequence shown here is derived from an EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is preliminary data.		
	DR		
	Hypothetical protein.		
	SQ		
	SEQUENCE 823 AA: 83110 MW;	CAC6C9027312BBBA CRC64;	
	Query Match	Score 105.5; DB 2;	Length 823;
	Best Local Similarity	59.1%; Pred. No. 0.47;	
	Matches 26; Conservative 7; Mismatches 8; Indels 3; Gaps 1;		
	KW		
	CC Preliminary data.		
	DR		
	EMBL; ARCU0100400; EAA54112.1; - ; Genomic_DNA.		
	RL		
	Submitted (OCT-2003) to the EMBL/GenBank/DDBJ databases.		
	CC		
	CC CAUTION: The sequence shown here is derived from an EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is preliminary data.		
	DR		
	Hypothetical protein.		
	SQ		
	SEQUENCE 823 AA: 83110 MW;	CAC6C9027312BBBA CRC64;	
	Query Match	Score 105.5; DB 2;	Length 823;
	Best Local Similarity	59.1%; Pred. No. 0.47;	
	Matches 26; Conservative 7; Mismatches 8; Indels 3; Gaps 1;		
	KW		
	CC Preliminary data.		
	DR		
	EMBL; ARCU0100400; EAA54112.1; - ; Genomic_DNA.		
	RL		
	Submitted (OCT-2003) to the EMBL/GenBank/DDBJ databases.		
	CC		
	CC CAUTION: The sequence shown here is derived from an EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is preliminary data.		
	DR		
	Hypothetical protein.		
	SQ		
	SEQUENCE 823 AA: 83110 MW;	CAC6C9027312BBBA CRC64;	
	Query Match	Score 105.5; DB 2;	Length 823;
	Best Local Similarity	59.1%; Pred. No. 0.47;	
	Matches 26; Conservative 7; Mismatches 8; Indels 3; Gaps 1;		
	KW		
	CC Preliminary data.		
	DR		
	EMBL; ARCU0100400; EAA54112.1; - ; Genomic_DNA.		
	RL		
	Submitted (OCT-2003) to the EMBL/GenBank/DDBJ databases.		
	CC		
	CC CAUTION: The sequence shown here is derived from an EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is preliminary data.		
	DR		
	Hypothetical protein.		
	SQ		
	SEQUENCE 823 AA: 83110 MW;	CAC6C9027312BBBA CRC64;	
	Query Match	Score 105.5; DB 2;	Length 823;
	Best Local Similarity	59.1%; Pred. No. 0.47;	
	Matches 26; Conservative 7; Mismatches 8; Indels 3; Gaps 1;		
	KW		
	CC Preliminary data.		
	DR		
	EMBL; ARCU0100400; EAA54112.1; - ; Genomic_DNA.		
	RL		
	Submitted (OCT-2003) to the EMBL/GenBank/DDBJ databases.		
	CC		
	CC CAUTION: The sequence shown here is derived from an EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is preliminary data.		
	DR		
	Hypothetical protein.		
	SQ		
	SEQUENCE 823 AA: 83110 MW;	CAC6C9027312BBBA CRC64;	
	Query Match	Score 105.5; DB 2;	Length 823;
	Best Local Similarity	59.1%; Pred. No. 0.47;	
	Matches 26; Conservative 7; Mismatches 8; Indels 3; Gaps 1;		
	KW		
	CC Preliminary data.		
	DR		
	EMBL; ARCU0100400; EAA54112.1; - ; Genomic_DNA.		
	RL		
	Submitted (OCT-2003) to the EMBL/GenBank/DDBJ databases.		
	CC		
	CC CAUTION: The sequence shown here is derived from an EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is preliminary data.		
	DR		
	Hypothetical protein.		
	SQ		
	SEQUENCE 823 AA: 83110 MW;	CAC6C9027312BBBA CRC64;	
	Query Match	Score 105.5; DB 2;	Length 823;
	Best Local Similarity	59.1%; Pred. No. 0.47;	
	Matches 26; Conservative 7; Mismatches 8; Indels 3; Gaps 1;		
	KW		
	CC Preliminary data.		
	DR		
	EMBL; ARCU0100400; EAA54112.1; - ; Genomic_DNA.		
	RL		
	Submitted (OCT-2003) to the EMBL/GenBank/DDBJ databases.		
	CC		
	CC CAUTION: The sequence shown here is derived from an EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is preliminary data.		
	DR		
	Hypothetical protein.		
	SQ		
	SEQUENCE 823 AA: 83		

ALIGNMENTS

Searched:	283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters:	283416

minimum DB seq length: 0
RESULT 1
HSURIP

Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Citation: Gonatida - sea urchin (*Terechinus angulosus*)
nstone n., Parchinius angulosus (argulus urchin)
C;Species: Parchinius angulosus (argulus urchin)
C;Date: 31-Mar-1980 #sequence 31-Mar-1980 #text_change 09-Jul-2004
C;Accession: A91090 #seqid 1 A91091 A02586

R;Strickland, W.N.; Strickland, M.; de Groot, P.C.; von Holt, C.; Wittmann-Liebold, R.; Biochem., 104, 559-566, 1980
Eur. J. Biochem.

Ali: The primary structure of histone H1 from sperm of the sea urchin *Ritterellia*.
Amino cyanogen bromide peptides.
A: Reference number: A91090; PMID:80156831; PMID:6767609

A;Contents: sequence of residues 1-84
A;Accession: A91090
A;Molecule type: protein

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, A;Residues: 1-248 <STR> A;Cross-references: UNIPROT:P02256; UNIPARC:UPI000012BF3D.

and is derived by analysis of the total score distribution.
SUMMARIES
A;Title: The primary structure of histone H1 from sperm of the sea urchin Pa-
Eur. J. Biochem. 104, 567-578, 1980
R;Strickland, W.N.; Strickland, M.; Brandt, W.F.; von Holt, C.; Lehmann, A.;

A;Reference number: A91091; MUID:801566832; PMID:7363905
A;Accession: A91091
A;Molecule type: protein
D;Denomination:

AMINO ACID	SCORE	MATCH	LENGTH	DESCRIPTION
1	103	49.3	248	HSUR1P histone H1, gonadal

A: Note: 144-Arg was also found
 C: Superfamily: histone H1
 C: Keywords: DNA binding; nucleosome; sperm

	best Local Similarity	%	Pre. NO.	U-0.91;	Indels	Gaps
/	S25194	43.3				
8	T18513	166	2			
9	JC7219	229	2			
					4;	Mismatches
					13;	
					25;	Conservative

		Qy	Db
10	90.5	A86315	F2H15.19 protein -
11	90	383	histone H2B (clone
12	90	2	homoeetical
		2	146
		2	2

12	20	43.1	529	hypothetical prote
13	90	43.1	529	hypothetical prote
14	89	42.6	208	histone H1.1 - Cae

cgcrr-4 protein - C hypothetical prote hypothetical prote hypothetical protein T20N10.250 - Arabidopsis thaliana

```

C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004
C;Version: m0172

```

Accession #: U42153
 R.D'Angelo, M.; Vezzi, A.; Modesto, D.; Pigazzi, M.; Valle, G.; Mewes, H.W.; submitted to the Protein Sequence Database, April 2000

A; Reference number: Z25017
 A; Accession: T49173
 A; Status: Preliminary
 ribosomal protein
 histone H2B (clone
 histone H1-H1 - Vo

SAR DNA-binding protein
probable RNA-binding protein

ALIGNMENTS

```

RESULT 1
AAB13783
ID AAB13783 standard; peptide; 45 AA.
XX
AC
AAB13783;
XX
DT 10-NOV-2000 (first entry)
XX
DE Soluble tandem pEA/ pk peptide conjugate.
XX
KW pk peptide; cytotoxic; vaccine; cytotoxic T cell; CTL; immunotherapy;
KW major histocompatibility complex class I; MHC class I; antigen; tumour;
KW prostate; breast; multiple myeloma; DEA peptide.
KW

```

No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.

סימן

result	No.	Score	Query	Match	Length	DB	ID	Description
1	209	100.0	AAB13783	45	3	Aab13783	Soluble t	Aab13783 Soluble t
2	119	56.9	AAB23585	36	3	Aab23585	Ask21 lin	Aab23585 Synthetic
3	119	56.9	ADH89966	211	8	Adbh9966	Synthetic	Adbh9966 Synthetic
4	119	56.9	AAB23591	630	3	Aab23591	Modified	Aab23591 Modified
5	119	56.9	AAB23593	640	3	Aab23593	Drosophila	Aab23593 Drosophila
6	118	56.5	ABBT1574	2451	4	Abb1574	Human pol	Aao01368 Human pol
7	115	55.3	AAQ01368	123	4	Aao01368	Human pol	Aao01368 Human pol
8	115	55.0	AAQ03703	113	4	Aao03703	Human pol	Aao03703 Human pol
9	113	54.1	ABP6694	122	5	Abp6694	Human bre	Abp6694 Human bre
10	111	53.1	AAU19275	63	4	Aau19275	Novel hum	Aau19275 Novel hum
11	111	53.1	ABG32696	63	5	Abg32696	Human DNA	Abg32696 Human DNA
12	111	53.1	ADC25413	63	7	Adc25413	Human ext	Adc25413 Human ext
13	110	52.6	ABG27118	141	4	Abg27118	Novel hum	Abg27118 Novel hum
14	109	52.2	AAB13781	25	3	Aab13781	Soluble P	Aab13781 Soluble P
15	109	52.2	AAV98495	59	3	Aay98495	Nuclear 1	Aay98495 Nuclear 1
16	109	52.2	AAV9040	59	3	Aay9040	Nuclear 1	Aay9040 Nuclear 1
17	109	52.2	AAB5848	59	4	Aab5848	Nucleic a	Aab5848 Nucleic a
18	109	52.2	AAU01285	59	4	Aau01285	Nuclear 1	Aau01285 Nuclear 1
19	109	52.2	AAO03278	74	4	Aao03278	Human pol	Aao03278 Human pol
20	109	52.2	AAO08943	75	4	Aao08943	Human pol	Aao08943 Human pol
21	109	52.2	AAU02477	112	4	Aau02477	Human pol	Aau02477 Human pol
22	109	52.2	AAU8240	110	4	Aau8240	Novel hum	Aau8240 Novel hum
23	109	52.2	ABG92661	150	5	Abg92661	Human DNA	Abg92661 Human DNA

XX 14-DEC-1999; 99WO-US029724.
XX PF
XX PR 14-DEC-1998; 98US-0112324P.
XX PA (DEND-) DENDRON CORP.
XX PI Laus R, Hakim I, Vidovic D;
XX DR WPI; 2000-442365/38.
XX PT Antigens modified by the covalent addition of a peptide that facilitates
PT entry into antigen presenting cells, useful for producing compositions
PT for immunizing against tumors and pathogens.
XX
PS Claim 2; Page 26; 34pp; English.
XX
CC The present invention relates to compositions of modified soluble protein
CC antigens capable of eliciting an enhanced in vivo cytotoxic T cell (CTL)
CC response i.e. a major histocompatibility complex (MHC) class I molecule
CC response. The protein antigen is modified by the covalent addition of a
CC peptide sequence which facilitate entry of the antigen into antigen
CC presenting cells (APCs). The present sequence is one such peptide
CC sequence which can be used to modify the soluble antigens. The present
CC sequence is tandem PEA/ PK peptide conjugate. The modified antigen
CC composition may be used for immunising against, or treating a tumour e.g.
CC prostate and breast carcinoma or multiple myeloma, or pathogen in mammals

PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 14-AUG-2000; 2000US-0225759P.
 PR 18-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226631P.
 PR 22-AUG-2000; 2000US-0226868P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-02709P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229290P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 06-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0229517P.
 PR 06-SEP-2000; 2000US-022987P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0229345P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 12-SEP-2000; 2000US-0231242P.
 PR 14-SEP-2000; 2000US-0231244P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233053P.
 PR 14-SEP-2000; 2000US-0233054P.
 PR 14-SEP-2000; 2000US-0233055P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 25-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234938P.
 PR 26-SEP-2000; 2000US-0235084P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 29-SEP-2000; 2000US-0235836P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236569P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0237584P.
 PR 13-OCT-2000; 2000US-0237585P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246533P.
 PR 08-NOV-2000; 2000US-0246534P.

PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0245526P.
 PR 08-NOV-2000; 2000US-0245527P.
 PR 08-NOV-2000; 2000US-0245528P.
 PR 08-NOV-2000; 2000US-0245532P.
 PR 08-NOV-2000; 2000US-0245609P.
 PR 08-NOV-2000; 2000US-024610P.
 PR 08-NOV-2000; 2000US-024611P.
 PR 08-NOV-2000; 2000US-024613P.
 PR 08-NOV-2000; 2000US-024613P.
 PR 17-NOV-2000; 2000US-024910P.
 PR 17-NOV-2000; 2000US-024910P.
 PR 17-NOV-2000; 2000US-024910P.
 PR 17-NOV-2000; 2000US-024911P.
 PR 17-NOV-2000; 2000US-024911P.
 PR 17-NOV-2000; 2000US-024912P.
 PR 17-NOV-2000; 2000US-024913P.
 PR 17-NOV-2000; 2000US-024914P.
 PR 17-NOV-2000; 2000US-024914P.
 PR 17-NOV-2000; 2000US-024915P.
 PR 17-NOV-2000; 2000US-024915P.
 PR 17-NOV-2000; 2000US-024916P.
 PR 17-NOV-2000; 2000US-024917P.
 PR 17-NOV-2000; 2000US-024917P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-024924P.
 PR 17-NOV-2000; 2000US-024924P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251188P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-025156P.
 PR 08-DEC-2000; 2000US-025168P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-025189P.
 PR 08-DEC-2000; 2000US-025190P.
 PR 11-DEC-2000; 2000US-025497P.
 PR 05-JAN-2001; 2001US-0255678P.
 PR 17-JAN-2001; 2001US-00764846.

(HUMA-) HUMAN GENOME SCI INC.
 XX PA
 PI Rosen CA, Ruben SM, Barash SC;
 XX WPI; 2003-605749/57.
 DR N-PSBB; ADC25285.

XX PT New DNA-binding proteins and gene encoding them, useful for diagnosing, treating and/or preventing e.g. neurological, inflammatory, infectious, cardiovascular, autoimmune, respiratory, neoplastic or digestive diseases.

XX PS Claim 11; SEQ ID NO 260; 226pp; English.

XX PT The invention relates to an isolated nucleic acid molecule (cDNA) encoding a human extracellular matrix protein, representing one of 161 novel genes. Also included are recombinant vectors, host cells (expressing the protein), the extracellular matrix proteins (including their fragments and homologues), an isolated antibody that binds specifically to the protein, diagnosing a pathological condition or susceptibility to a pathological condition (comprising determining the presence or amount of expression of the protein in a biological sample and diagnosing a condition based on the presence or amount of expression of the protein). Preventing, treating or ameliorating a medical condition by administering the nucleic acid or protein to a mammalian subject, identifying a binding partner to the protein, the gene corresponding to the cDNA sequence, and identifying an

